Cancer Causes and Control **13**: 159–168, 2002.

A case–control study of tobacco use and other non-occupational risk factors for t(14;18) subtypes of non-Hodgkin's lymphoma (United States)

J.C. Schroeder[1,2,*], A.F. Olshan[2], R. Baric[2], G.A. Dent[3], C.R. Weinberg[4], B. Yount[2], J.R. Cerhan[5], C.F. Lynch[6], L.M. Schuman[7], P.E. Tolbert[8], N. Rothman[9], K.P. Cantor[9] & A. Blair[9]

[1]*Epidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA;* [2]*Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, 2103C McGavran-Greenberg Hall, CB 7435, Chapel Hill, NC, USA; Ph.: (919) 966-7446; Fax: (919) 966-2089;* [3]*Department of Pathology and Laboratory Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA;* [4]*Biostatistics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA.* [5]*Health Sciences Research, Mayo Clinic, Rochester, MN, USA;* [6]*Department of Epidemiology, University of Iowa, Iowa City, IA, USA;* [7]*Department of Epidemiology, University of Minnesota, Minneapolis, MN, USA;* [8]*Rollins School of Public Health, Emory University, Atlanta, GA, USA;* [9]*Occupational Epidemiology Branch, National Cancer Institute, Bethesda, MD, USA (*Author for correspondence)*

Received 13 February 2001; revised version accepted 18 October 2001

Key words: bcl-2, epidemiology, lymphoma, molecular genes, non-Hodgkin, tobacco, translocation (genetics).

Abstract

Objective: Non-Hodgkin's lymphoma (NHL) encompasses diverse subtypes, and analyzing NHL as a single outcome may mask associations. In a new approach we evaluated associations with subtypes defined by the t(14;18) translocation, reasoning that cases within these subtypes would have more common risk factors than all NHL combined.

Methods: Archival biopsies from cases in a population-based NHL study were assayed for t(14;18) using polymerase chain reaction amplification. Exposures in 68 t(14;18)-positive and 114-negative cases were compared with 1245 controls. The expectation–maximization algorithm was used to fit polytomous regression models based on all available information, including data from 440 unclassified cases.

Results: Family history of hemolymphatic cancer was associated with t(14;18)-negative NHL (odds ratio (OR) 2.4, 95% confidence interval (CI) 1.4–3.9), but not t(14;18)-positive NHL. Cigarette smoking was weakly associated with t(14;18)-positive NHL (OR 1.7, CI 0.9–3.3), but ORs decreased as smoking increased. Chewing tobacco was associated with t(14;18)-positive NHL, particularly when used before age 18 (OR 2.5, CI 1.0–6.0, 13 exposed cases). Odds ratios for both case-subtypes were doubled among hair-dye users.

Conclusions: Cigarette smoking was not clearly associated with t(14;18)-positive NHL. Family history may be a marker for factors that act specifically through t(14;18)-negative pathogenic mechanisms.

Introduction

Cases of non-Hodgkin's lymphoma (NHL) vary in histologic appearance and clinical behavior, and it is widely accepted that their etiologic determinants probably vary also [1–5]. If this is true we may fail to identify risk factors that are relevant only to subsets of NHL cases when we estimate relative risks for all NHL combined [6]. Recent studies support the hypothesis that risk factors may be more strongly associated with NHL subtypes than with NHL in the aggregate [7–12], but the case-subtype definitions used in these studies have rarely coincided, and all were based on classification schemes designed to group cases according to clinical characteristics, rather than etiologic determinants. Unfortunately it is not at all obvious how to define etiologically relevant subtypes without prior knowledge of the etiologic factors involved.

In a new approach we subtyped NHL cases based on the t(14;18) translocation, the most common somatic

mutation in NHL. The t(14;18) translocation joins the immunoglobulin heavy chain gene (*IgH*) on chromosome 14 with the *bcl-2* gene on chromosome 18, resulting in constant production of bcl-2 protein, a key inhibitor of apoptosis [13–16]. By itself t(14;18) is not sufficient to cause neoplastic transformation, but *bcl-2* activation may increase the risk of NHL by preventing cell death and allowing subsequent oncogenic mutations to develop and persist [1, 5]. Although subtype-specific associations could arise through different causal mechanisms (*e.g.* risk factors could increase the incidence of t(14;18), the prevalence of t(14;18)-positive lymphocytes, or the likelihood that t(14;18)-positive lymphocytes will progress to t(14;18)-positive NHL), each mechanism would share t(14;18) as a common component. Thus, we reasoned that t(14;18)-subtyping would increase the pathogenic specificity of our outcome groups relative to all NHL combined, and this in turn might lead to a relative increase in etiologic specificity. In this report we focus on relations between t(14;18)-positive and -negative NHL subtypes and several factors that have been inconsistently associated with NHL as a whole, including tobacco use, hair-dye use, and a history of hemolymphatic cancer in blood relatives.

Materials and methods

Cases

Study participants were originally enrolled in the National Cancer Institute's Factors Affecting Rural Men (FARM) study, a population-based case–control study of risk factors for incident hemolymphatic cancers among white men in Iowa and Minnesota [17, 18]. The State Health Registry of Iowa, a member of the National Cancer Institute's Surveillance and Epidemiology End Results (SEER) program, identified incident NHL cases diagnosed in Iowa between March 1981 and October 1983. The Environmental and Occupational Health Division of the University of Minnesota School of Public Health ascertained Minnesota cases diagnosed between October 1980 and September 1982, through active surveillance of state hospital and pathology records. The study was restricted to white men age 30 and older and, in Minnesota, to cases residing outside of major metropolitan areas (metropolitan Minneapolis, St Paul, Duluth, Rochester). Seven-hundred eighty cases of NHL were ascertained for the FARM study, 694 (89%) cases (or next-of-kin proxies) were interviewed, and 622 (90%) of the interviewed cases were confirmed to be NHL by pathology review [17].

Case-subtypes

To determine translocation status we assayed archival tumor tissue from a subset of FARM study cases. The State Health Registry of Iowa and the Environmental and Occupational Health Division of the University of Minnesota School of Public Health requested paraffin-embedded tumor blocks from all FARM study NHL cases. Block retrieval began in early 1997, and took about 18 months to complete. Archival samples were successfully retrieved for 40% (248) of the 622 FARM study NHL cases. About half the non-retrieved cases were destroyed or lost, while others were held at institutions that declined to participate in the current study. An experienced histopathologist (G.D.) classified cases according to a modified form of the Revised European–American Lymphoma classification system (REAL), based on review of newly cut sections and information contained in the original pathology report, without knowledge of exposure or translocation status [19]. Translocation assays were performed at the University of North Carolina School of Public Health, under the supervision of one of the study collaborators (R.B.). Laboratory personnel did not know the exposure status or histologic classification of individual cases.

Translocation status was determined using a polymerase chain reaction (PCR) assay designed to amplify DNA spanning the *bcl-2* and *IgH* breakpoints most often joined by the translocation. Each reaction included a single primer corresponding to conserved segments in the J (joining) region of the *IgH* gene on chromosome 14 (J_H consensus primer: 5′-ACCTGAGGAGACGGTGACC-3′) and two primers corresponding to *bcl-2* segments in the major breakpoint region (MBR) of chromosome 18. One *bcl-2* primer was similar to the MBR primer published by Crescenzi *et al.* (MBR1: 5′-GAGAGTTGCTTTACGTGGCCTG-3′) [20], the second corresponded to a segment of *bcl-2* 360 base pairs upstream from the MBR (MBR2: 5′-CGCTTGACTCCTTTACGTGCTG-3′). These primers would not amplify DNA if t(14;18) were absent, if translocation breakpoints fell outside of the primers, or if DNA degradation disrupted breakpoint amplification targets. DNA integrity was confirmed based on amplification of a 175-base pair control segment of the normal β-globin gene [21].

DNA was extracted from paraffin sections using a phenol:chloroform extraction procedure. Following PCR, amplified DNA was isolated, denatured, and transferred to a nylon membrane, using gel electrophoresis and Southern blotting. The β-globin amplification was confirmed based on visualization of a 175-base pair fragment following electrophoresis. Amplification of t(14;18) was confirmed using radiolabeled probes de-

signed to anneal to DNA adjacent to *bcl-2* breakpoints (MBR1 probe: 5′-CAACACAGACCCACCCAGAGC-3′; MBR2 probe: 5′-GATGGCTTTGCTGAGAGGT-TTG-3′). A subset of cases was probed for J_H consensus segments in *IgH* (J_H probe: 5′-GGGTBCCWTGGCCC-CAG-3′, B = GCT, W = AT). All β-globin-positive PCR products that were negative for t(14;18) on the first round of PCR (n = 125) were subjected to a nested PCR assay, using interior primers corresponding to the probes listed above. Eleven cases were classified as t(14;18)-positive based on the results of the nested assay.

Sequenced t(14;18) breakpoint amplification products were less than 150 base pairs in 8/20 samples, which suggests that in some cases DNA integrity may have been sufficient for t(14;18) amplification, but not amplification of the 175-base pair β-globin control. Consequently, we classified 10 β-globin negative cases that amplified a t(14:18) breakpoint in the first (non-nested) PCR assay as t(14;18)-positive. Samples that failed to amplify either product were considered unclassified (66 cases, 27% of the retrieved blocks).

Strict precautions to avoid sample contamination were taken in all steps of the t(14;18) assay. Amplification of uniform breakpoint products would strongly suggest sample contamination, since t(14;18) breakpoints are generated and joined in a process that is highly variable [15]. Therefore, we monitored amplification products for size variation, and confirmed amplification of unique DNA segments in a sample (20/68) of t(14;18)-positive cases by cloning and automated sequencing.

Controls

All FARM study controls were included in this analysis. They included 1245 white men age 30 or older without hemolymphatic cancer, frequency-matched to cases based on state, vital status, and age within 5-year age groups. Men from metropolitan areas of Minnesota were not eligible [17]. Controls for living cases were recruited from eligible men identified by random-digit dialing (men under age 65, 77% response), and Health Care Financing Administration Medicare Files (men 65 years and older, 79% response). Controls for deceased cases were identified using state death certificate files, and 77% of eligible next-of-kin agreed to participate.

Exposure assessment

Structured in-person interviews were administered between August 1981 and May 1984 [17, 18]. Participants were asked for detailed information about education, marital status, ancestral origins, religion, residential history, tobacco use, coffee and alcohol consumption, pet dogs and cats, hobbies, medical history, agricultural exposures, occupational history, and occupational or other exposure to potential NHL risk factors that occurred at least once a month for a year or more.

Tobacco and individual tobacco products (cigarettes, pipe tobacco, cigars, chewing tobacco, snuff) were classified as ever- or never-used (based on daily use for at least 3 months), and according to age of first use (at or before age 18, after age 18). Cigarette smoking was classified according to exclusive or combined use of cigarettes with other tobacco products, and was categorized according to average cigarettes per day, and years or pack-years of cigarettes up to 2 years prior to the date of diagnosis (cases) or interview (controls).

Participants were asked whether blood relatives had been diagnosed with leukemia, lymphoma, or multiple myeloma [17, 18], and were categorized based on disease in first-degree relatives only (parents, children, siblings), or disease in first- or second-degree relatives (grandparents, grandchildren, aunts, uncles, nieces, nephews, and half-siblings). Hair-dye use was based on ever use of hair tints or any hair-coloring products, use at least once a month for a year or more, or occupational exposure on any job held 1 year or longer.

Data analysis

Analyses were run using Stata, Release 6.0 [22]. Associations between exposures and all FARM study NHL cases compared with controls were evaluated using unconditional logistic regression models adjusted for frequency-matching factors (age, state, and vital status) [17]. Logistic regression models also were used to estimate case–case odds ratios for t(14;18)-positive compared with -negative NHL cases.

Missing data were a problem for comparisons of NHL subtypes with controls, since 70% of cases were not classifiable, usually because archival tumor blocks were not available. Missing case-subtype information was obviously related to being a case. If missing case-subtype data also were related to one or more covariates, then the missing data could be informative, and simple exclusion of missing cases from the polytomous regression could bias estimation of corresponding odds ratios [23]. Instead, we used an approach to model fitting that maximized the likelihood based on all available data, including data from unclassified cases, by applying a statistical missing-data technique called the expectation–maximization (EM) algorithm [24, 25]. In brief, we would begin by modeling the complete case data (controls and case observations with t(14;18) case-subtype information) using unconditional polytomous

logistic regression [25, 26]. Model results were used to derive the estimated probability of each case-subtype among cases within each covariate stratum, and missing case-subtype observations were apportioned to t(14;18)-positive and -negative case-subtypes according to these expectations. These pseudo-data were modeled using polytomous logistic regression in a maximization step, and new estimated probabilities were used to create new pseudo-data in a subsequent expectation step. Iterations between maximization and expectation steps continued in this manner until models converged. Odds ratios were based on estimates from the final model, and variances were calculated based on the observed-data likelihood. EM method results should be valid as long as missing case-subtype data are unrelated to case-subtypes within covariate strata, *i.e.* when the subtype data are missing at random (MAR) [24]. Using a simulation study designed to approximate the conditions of this analysis, we confirmed that for case-subtype data missing at random the EM method could improve precision and prevent bias caused by ignoring unclassified cases [25].

Initial models included state and age (coded using upper and lower tail-restricted quadratic splines) [27], vital status, all other primary exposures in this analysis, and either a single covariate indicating work as a farmer for at least six months after age 18, or two covariates indicating exposure to agricultural factors associated with t(14;18)-positive and -negative NHL in this population [28]. Model covariates were retained if elimination resulted in more than a 15% change in the coefficient of interest. There was little evidence of confounding based on this criterion, and final models included only state and age, unless otherwise indicated.

Joint effects of *a-priori* interest were modeled using indicator variables coded according to separate or joint exposure to pairs of dichotomous covariates, with those having neither exposure acting as a common referent group. Observed joint odds ratios were compared with those predicted for average additive or multiplicative effects, and interaction contrast ratios were estimated [29] in order to quantify statistical interactions on the additive scale.

Results

In total, 63% (114) of the successfully assayed cases were t(14;18)-negative, and 37% (68) were t(14;18)-positive. Proportions of assayed cases within histologic subtypes were similar to those for FARM study cases as a whole (Table 1). Proportions of t(14;18)-positive and -negative cases within REAL and Working Formulation subtypes (from the original FARM study pathology review) were compatible with previous reports [19, 30, 31].

Table 1. Distribution of FARM study NHL cases, assayed cases, and t(14;18) subtypes within Working Formulation or modified Revised European–American Lymphoma (REAL) histologic subtypes (figures in parentheses are percentages)

	Histologic subtype		t(14;18) status	
	All cases	Assayed	Negative	Positive
Working formulation[a]				
Small lymphocytic	85 (14)	29 (16)	22	7 (24)
Follicular	195 (31)	65 (36)	37	28 (43)
Diffuse small, mixed	86 (14)	22 (12)	10	12 (55)
Diffuse large cell	113 (18)	29 (16)	22	7 (24)
Immunoblastic	33 (5)	6 (3)	3	3 (50)
Lymphoblastic	6 (1)	2 (1)	2	0 (0)
Small non-cleaved	18 (3)	10 (6)	8	2 (20)
Unclassified	86 (14)	19 (10)	10	9 (47)
REAL[b]				
Chronic lymphocytic	–	23 (13)	20	3 (13)
Follicular	–	54 (30)	30	24 (44)
Mantle cell	–	3 (2)	2	1 (53)
Marginal zone	–	2 (1)	1	1 (50)
Diffuse large cell	–	64 (35)	39	25 (39)
Burkitt's/Burkitt-like	–	14 (8)	10	4 (29)
Unclassified	–	22 (12)	12	10 (45)
Total	622 (100)	182 (100)	114	68 (37)

[a] Working formulation subtypes determined during FARM study pathology review.
[b] Modified REAL subtypes determined during review of newly cut sections prior to t(14;18) assay, based on morphology only.

Assayed cases were more likely than the original FARM study case group to have been diagnosed in Iowa (58% vs 47%), and to have completed their own interview (76% vs 69%), but the age distribution of assayed cases was comparable to the original study sample (Table 2). Odds ratios based on all FARM study cases suggested an increase in the risk of NHL associated with hair-dye use, and with a family history of hemolymphatic cancer, as previously reported [32, 33]. Almost 80% of study participants used tobacco products.

The relative risk estimate for t(14;18)-positive NHL compared with controls was increased somewhat in association with use of any tobacco (OR 1.5, CI 0.8–2.8), and with the exclusive use of cigarettes (OR 1.7, CI 0.9–3.3) (Table 3). However, the association between cigarette use and t(14;18) NHL was limited to the lowest categories of pack-years, cigarettes smoked per day, and years of cigarette smoking. Smokers who started smoking after age 18 had a greater risk of t(14;18)-positive NHL (OR 1.6, CI 0.9–2.8) than those who began at or before age 18 (OR 1.1, CI 0.6–2.0). Odds ratios for t(14;18)-negative NHL and cigarette smoking were close to the null for all exposure indices. Use of chewing tobacco was associated with t(14;18)-positive NHL (OR 1.7, CI 0.9–3.1), particularly when use began at age 18 or younger (OR 2.5, CI 1.0–6.0). Cigar smoking was weakly associated with t(14;18)-negative NHL, and use was uncommon prior to age 19.

Hair-dye use was associated with a two-fold increase in relative risk for both t(14;18)-positive and -negative case-subtypes (Table 4). Family history of hemolymphatic

Table 2. Distribution of matching factors and selected characteristics among controls, FARM study NHL cases, and cases assayed for t(14;18), and odds ratios[a] for FARM study NHL cases compared with controls

	Percentage controls	Percentage all NHL	Percentage assayed cases	All FARM cases: controls	
	(n = 1245)	(n = 622)	(n = 182)	OR	95% CI
State of Origin					
Iowa	48	47	58	–	
Minnesota	52	53	42	–	
Vital status					
Alive	66	69	76	–	
Deceased	34	31	24	–	
Age (years)					
30–55	20	26	24	–	
>55–65	23	23	21	–	
>65–75	28	26	30	–	
>75	29	25	24	–	
Marital status					
Married/widowed	91	90	88	1.0	
Divorced/separated	4	4	6	1.1	0.7–1.2
Never married	6	5	6	1.0	0.6–1.5
Family history[b]					
No	94	90	93	1.0	
Yes	4	8	5	2.0	1.3–2.9
Don't know	2	2	2	–	–
Farmer	56	57	60	1.1	0.9–1.4
Any hair-dye use	5	9	14	2.0	1.3–2.9
Education ⩽12 years.	71	70	70	1.0	0.8–1.3
Drank alcohol	55	55	54	0.9	0.8–1.1
Used any tobacco	77	81	81	1.3	1.0–1.7
Cigarettes	66	70	68	1.1	0.9–1.4
Pipe	24	25	26	1.1	0.9–1.4
Cigar	13	15	14	1.2	0.9–1.6
Chewing tobacco	9	11	11	1.3	0.9–1.8
Snuff	11	10	11	1.0	0.7–1.4

[a] Odds ratios from unconditional logistic regression models, adjusted for frequency-matching factors (state, vital status, age). Exposures without a specified referent category are dichotomous.

[b] History of hemolymphatic cancer in parent, sibling, or child.

Table 3. Adjusted[a] odds ratios for use of tobacco products among t(14;18)-positive or -negative NHL cases compared with controls, and for t(14;18)-positive cases compared with -negative cases

Exposure	No. of controls	t(14;18)-positive NHL: controls[b]			t(14;18)-negative NHL: controls[b]			t(14;18)-positive: t(14;18)-negative[c]	
		No.	OR	95% CI	No.	OR	95% CI	OR	95% CI
No tobacco use	286	10	1.0		24	1.0		1.0	
Used any tobacco	959	58	1.5	0.8–2.8	89	1.2	0.8–1.7	1.5	0.7–3.5
Tobacco by product[d]									
Cigarettes	825	50	1.3	0.8–2.2	73	1.0	0.7–1.4	1.4	0.7–2.9
Pipe	293	20	1.3	0.8–2.0	27	1.0	0.7–1.4	1.3	0.6–2.8
Cigar	167	7	0.7	0.4–1.5	18	1.4	0.9–2.1	0.6	0.2–1.5
Chewing tobacco	115	10	1.7	0.9–3.1	9	1.0	0.6–1.8	1.8	0.6–5.1
Snuff	141	7	1.0	0.5–2.0	12	0.9	0.6–1.6	0.9	0.3–2.8
No tobacco use	286	10	1.0		24	1.0		1.0	
Cigarettes only	454	30	1.7	0.9–3.3	38	1.1	0.7–1.7	1.9	0.8–4.6
Cigarettes and other	371	20	1.4	0.7–2.8	35	1.2	0.8–1.8	1.4	0.5–3.5
Other tobacco only	132	8	1.5	0.6–3.4	16	1.5	0.9–2.5	1.0	0.3–3.4
Cigarette pack-years									
0	423	18	1.0		40	1.0		1.0	
>0–20	219	16	1.6	0.9–3.0	19	1.1	0.7–1.7	2.0	0.8–4.9
>20–40	219	10	1.1	0.5–2.2	16	0.9	0.5–1.4	1.4	0.5–3.9
>40	326	18	1.1	0.6–2.0	33	1.0	0.7–1.5	1.2	0.5–2.7
Cigarettes per day									
0	418	18	1.0		40	1.0		1.0	
>0–10	176	14	1.8	1.0–3.5	12	0.9	0.5–1.5	2.6	1.0–6.8
>10–20	363	20	1.2	0.7–2.2	34	1.0	0.7–1.5	1.3	0.6–3.0
>20	259	13	1.0	0.5–2.0	26	1.1	0.7–1.6	1.0	0.4–2.5
Product and age started[d]									
Cigarettes									
>18 years old	332	22	1.6	0.9–2.8	26	0.9	0.6–1.4	1.9	0.8–4.5
≤18 years old	493	28	1.1	0.6–2.0	47	1.1	0.8–1.6	1.1	0.5–2.4
Chewing tobacco									
>18 years old	101	59	1.3	0.6–2.9	8	1.2	0.6–2.2	1.3	0.3–4.7
≤18 years old	32	13	2.5	1.0–6.0	16	1.0	0.3–3.0	3.3	0.7–16.2

[a] Adjusted for state and age (restricted quadratic splines), unless otherwise indicated.
[b] Estimates from EM polytomous regression models including controls and both case-subtypes.
[c] Estimates from unconditional logistic regression models restricted to assayed cases.
[d] Estimates from single model, relative to non-users of each product.

cancer was associated with t(14;18)-negative NHL (OR 2.4, CI 1.4–3.9 for history in a parent, sibling, or child), but not t(14;18)-positive NHL (OR 1.3, CI 0.5–3.3). Odds ratios for t(14;18)-positive NHL were about doubled for never-married men and for divorced or separated men, relative to men who were married or widowed, while corresponding odds ratios for t(14;18)-negative NHL were below the null.

Estimates of relative risks for t(14;18)-positive and -negative NHL in association with joint exposures (to different tobacco products, to tobacco products and other non-occupational exposures, or to a family history of hemolymphatic cancer and other non-occupational exposures) were unstable but consistent with additive effects (data not shown).

Discussion

These data fail to provide clear support for an association between smoking and t(14;18)-positive NHL, given the inverse dose–response pattern noted with increased duration or amount of smoking. Several previous studies, including the FARM study, did not find convincing evidence of an association between cigarette smoking and all NHL [7, 8, 10, 34], while others have reported weak positive associations [9, 35, 36]. Studies have also reported associations between smoking and follicular lymphoma, a histologic subtype with a relatively high proportion of t(14;18)-positive cases [8, 30, 37]. In addition, heavy smokers were reported to have more t(14;18)-positive lymphocytes in

Table 4. Adjusted[a] odds ratios for selected exposures among t(14;18)-positive or -negative NHL cases compared with controls, and for t(14;18)-positive cases compared with -negative cases

Exposure	No. of controls	t(14;18)-positive NHL: controls[b]			t(14;18)-negative NHL: Controls[b]			t(14;18)-positive: t(14;18)-negative[c]	
		No.	OR	95% CI	No.	OR	95% CI	OR	95% CI
State									
Iowa	603	41	1.0		64	1.0		1.0	
Minnesota	642	27	1.0	0.6–1.5	50	1.2	0.9–1.5	0.9	0.5–1.6
Age at diagnosis (years)									
30–55	253	18	1.0		28	1.0		1.0	
>55–65	283	12	0.7	0.3–1.3	26	0.9	0.6–1.4	0.7	0.3–1.8
>65–75	352	21	0.8	0.4–1.3	33	0.7	0.5–1.1	1.0	0.5–2.3
>75	357	17	0.8	0.4–1.4	27	0.7	0.5–1.1	1.0	0.4–2.5
Proxy interview	419	15	0.8	0.5–1.4	29	1.0	0.7–1.4	0.8	0.4–1.7
Any hair-dye use	58	9	1.8	0.9–3.7	17	2.1	1.3–3.4	0.9	0.4–2.1
Education ≤12 years	878	53	1.3	0.8–2.2	82	0.9	0.7–1.3	1.4	0.7–2.8
Drank alcohol weekly	686	38	1.0	0.6–1.6	60	0.9	0.6–1.2	1.1	0.6–2.1
Family history[d]									
First-degree	51	4	1.3	0.5–3.3	11	2.4	1.4–3.9	0.6	0.2–1.9
First- or second-degree	78	5	1.0	0.4–2.5	16	2.4	1.6–3.6	0.5	0.2–1.4
Married or widowed	1128	55	1.0		105	1.0		1.0	
Divorced or separated	47	6	2.0	0.9–4.1	5	0.7	0.3–1.6	2.7	0.7–9.9
Never married	70	7	1.9	1.0–3.7	4	0.5	0.2–1.2	3.2	0.9–11

[a] Adjusted for state and age (restricted quadratic splines).
[b] Estimates from EM polytomous regression models including controls and both case-subtypes.
[c] Estimates from unconditional logistic regression models restricted to assayed cases.
[d] Family history of hemolymphatic cancer. First-degree relatives: parents, children, siblings; second-degree: aunts/uncles, grandparents/children, nieces/nephews.

their peripheral blood than did non-smokers [38]. Matching on vital status may bias relative risk estimates toward the null for smoking and other factors associated with premature mortality, since these exposures may be over-represented among deceased controls [39, 40]. However, relative risk estimates generated with deceased controls or all deceased participants excluded were similar to estimates for the complete study sample (data not shown).

Based on a small number of exposed cases, we found evidence of an association between chewing tobacco and t(14;18)-positive NHL, particularly among those who began use at an early age. This preliminary finding may be worthy of further investigation, since recent estimates indicate that 14% of male high-school students in the United States use smokeless tobacco [41].

Hair-dye use has been associated with NHL in men [33], and women [42, 43], though other studies have reported negative findings [44, 45]. Studies that have evaluated NHL subtype-specific associations have noted stronger relations with follicular lymphoma than those reported for NHL in the aggregate [33, 42, 44]. In this study data did not support a subtype-specific association between hair dye use and t(14;18)-positive NHL.

Results support a subtype-specific association between family history of hemolymphatic cancer and t(14;18)-negative NHL, based on a small number of cases. A previous analysis of FARM study data reported associations of a similar magnitude between hemolymphatic cancer in parents and small lymphocytic lymphoma (OR 2.3, CI 0.7–7.2, 76% t(14;18)-negative in our study) and "other" subtypes (OR 2.8, CI 1.1–6.5, 62% t(14;18)-negative in our study) [32]. Family history of hemolymphatic cancer may be a marker for subtle forms of immune dysfunction [46], and NHL cases associated with immune suppression (due to congenital immune deficiencies [47], HIV [48], or therapeutic immune suppression [49]) also tend to occur in histologic subtypes that are predominantly t(14;18)-negative. In theory, subtype-specific associations with t(14;18)-negative NHL could occur if a factor competed with t(14;18) for causal mechanisms involving bcl-2. Epstein–Barr virus (EBV) infection, a cofactor in many lymphomas among immune-suppressed patients, can induce expression of the *bcl-2* gene in EBV-infected cells [50]. Thus it is possible that EBV (and immune suppression) might act specifically through t(14;18)-negative mechanisms. We are currently investigating whether bcl-2

protein is increased in tumors from t(14;18)-negative cases with a positive family history relative to other t(14;18)-negative cases in this study.

Never-married and divorced or separated men were more likely to have t(14;18)-positive NHL than were married or widowed men, but were less likely to have t(14;18)-negative NHL. These estimates did not appear to be confounded or modified by tobacco or alcohol use. Evidence for [11, 51] and against [52] an association between never-married or unmarried status and NHL has been reported. Marital status may be a marker for potentially high-risk behaviors related to sexual activity or recreational drug use [53], but intermediate factors other than AIDS would have to account for such a relation in the FARM study population [54, 55].

It is difficult to conceive of a causal mechanism that would account for the inverse association between marital status and t(14;18)-negative NHL. Discordant subtype-specific associations could result if t(14;18) developed in NHL clones that were already fully transformed, but t(14;18) is probably generated prior to neoplastic transformation in most, if not all, cases [1]. False-positive PCR results due to sample contamination could lead to discordant associations, but we believe that contamination was an unlikely cause of positive results in our study, and estimates were essentially unchanged when we excluded the most suspect t(14;18)-positive cases (*i.e.* those identified using nested-PCR, or t(14;18)-positive cases that were β-globin negative).

In previous studies PCR assays using primers designed to amplify both the MBR and the minor cluster region (mcr) failed to detect about 25% of t(14;18)-positive cases identified using cytogenetic analysis [56, 57]. We did not attempt to amplify mcr breakpoints (which may account for 10–15% of t(14;18)-positive cases [56]), but did use a novel upstream MBR primer that allowed us to identify nine t(14;18)-positive cases with breakpoints outside the range of primers used by others. Previous PCR-based studies have limited their analyses to follicular lymphomas or other selected subtypes; thus a direct comparison with our analysis of unselected NHL cases is not possible. However, the proportion of t(14;18)-positive cases among our follicular lymphomas (44%) was comparable to several previous reports [57–60], though somewhat lower than the proportion derived from the combined results of published studies (54% (201/375) t(14;18)-positive) [21, 56–61]. Disagreement among studies may indicate differences in assay sensitivity, but could also reflect population-specific variation in the pathogenic mechanisms that produce follicular lymphomas [31, 62, 63]. In fact, there is considerable variation in the proportion of t(14;18)-positive follicular lymphomas detected by cytogenetic analysis, the ''gold standard'' assay for t(14;18) (*e.g.* 50% of follicular lymphomas were t(14;18)-positive in a study reported by Levine *et al.* [64], in contrast to 85% reported by Yunis *et al.* [65]). We assessed the potential impact of false-negative results by recalculating estimates assuming 80% sensitivity and 100% specificity, and noted little change in t(14;18)-negative NHL:control comparisons. Odds ratios for t(14;18)-positive NHL were unchanged, as long as misclassification was non-differential with regard to exposure.

Although associations were evaluated for confounding by a variety of factors, bias due to uncontrolled confounding cannot be ruled out. Proxy respondents for deceased participants were often unable to provide information with regard to specific exposures [66], but vital status was not related to either case-subtype in our study, and misclassification of dichotomous exposures by proxy respondents should probably have led to bias toward the null [67]. However, matching on vital status may have biased relative risk estimates for factors associated with premature mortality, as previously noted [39, 40].

The extent of missing case-subtype data was a major limitation of this study. We used the EM method of model fitting in an attempt to correct bias due to incomplete case-subtype ascertainment, based on the assumption that the proportion of unclassified (missing) cases within covariate strata was the same in both case-subtype groups [24, 25]. We believe that this assumption was met because it was unlikely that translocation status would have been associated with the success of block retrieval; because clinical characteristics (grade, histologic subtype, tumor origin) were unrelated to block availability or the success of β-globin amplification; and because vital status was not associated with translocation status among assayed cases. However, the validity of this assumption cannot be tested. More importantly, use of the EM method cannot make up for problems related to small sample size. Consequently, our estimates are unstable and our results should be interpreted with caution unless confirmed.

In conclusion, we used available data and archival samples to evaluate whether associations with non-occupational factors would be stronger for t(14;18)-positive or -negative NHL case-subtypes than for all NHL combined, based on the hypothesis that t(14;18) would mark a subset of cases which share more of their etiologic basis than NHL cases overall. We did not find clear evidence of a subtype-specific association between smoking and t(14;18)-positive NHL, which suggests that t(14;18) was unrelated to smoking, or that translocation status did not define a case-subtype with a discernible

relation to smoking. In contrast, we noted evidence of a subtype-specific association between family history of hemolymphatic cancer and t(14;18)-negative NHL, based on a small number of cases. If confirmed, this suggests that t(14;18) and family history (or rather, the biologic mechanism it represents) might be competing components in a common pathogenic mechanism involving bcl-2.

Acknowledgements

We thank Letia Dean, David Cowen, and Dr Lynn Dressler for assistance with sample preparation; Freda Selk and Sarah Waldemar for assistance with tumor block retrieval; and Dr Amy Sayle and Dr Mariana Stern for comments on the manuscript. This work was supported by NIH Grant R03CA71617, and by a Cancer Epidemiology Research Award from the UNC Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill. Dr Cerhan was supported by NIH Grant K07 CA64220. The FARM study was funded by the National Cancer Institute, contract N01-CP-11020.

References

1. Magrath I (1992) Molecular basis of lymphomagenesis. *Cancer Res* **52s**: 5529s.
2. Jaffe ES, Raffeld M, Medeiros L, Stetler-Stevenson M (1992) An overview of the classification of non-Hodgkin's lymphomas: an integration of morphological and phenotypical concepts. *Cancer Res* **52s**: 5447s–5452s.
3. Weisenburger D (1992) Pathological classification of non-Hodgkin's lymphoma for epidemiological studies. *Cancer Res* **52s**: 5456s.
4. Herrinton L (1998) Epidemiology of the Revised European–American Lymphoma classification subtypes. *Epidemiol Rev* **20**: 187–203.
5. Potter M (1992) Pathogenetic mechanisms in B-cell non-Hodgkin's lymphoma in humans. *Cancer Res* **52s**: 5522s–5528s.
6. Weiss N, Liff J (1983) Accounting for the multicausal nature of disease in the design and analysis of epidemiologic studies. *Am J Epidemiol* **117**: 14–18.
7. Brown L, Everett G, Gibson R, Burmeister L, Schuman L, Blair A (1992) Smoking and risk of non-Hodgkin's lymphoma and multiple myeloma. *Cancer Causes Control* **3**: 49–55.
8. Herrinton L, Friedman G (1998) Cigarette smoking and risk of non-Hodgkin's lymphoma subtypes. *Cancer Epidemiol Biomarkers Prev* **7**: 25–28.
9. Freedman D, Tolbert P, Coates R, Brann E, Kjeldsberg C (1998) Relation of cigarette smoking to non-Hodgkin's lymphoma among middle-aged men. *Am J Epidemiol* **148**: 833–841.
10. Zahm S, Weisenburger D, Holmes F, Cantor K, Blair A (1998) Tobacco and non-Hodgkin's lymphoma: combined analysis of three case–control studies (United States). *Cancer Causes Control* **8**: 159–166.
11. Tatham L, Tolbert P, Kjeldsberg C (1997) Occupational risk factors for subgroups of non-Hodgkin's lymphoma. *Epidemiology* **8**: 551–558.
12. Scherr PA, Hutchinson G, Neiman R (1992) Non-Hodgkin's lymphoma and occupational exposure. *Cancer Res* **52s**: 5503s–5509s.
13. Cleary M, Smith S, Sklar J (1986) Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin transcript resulting from the t(14;18) translocation. *Cell* **47**: 19–28.
14. Tsujimoto Y, Croce C (1986) Analysis of the structure, transcripts, and protein products of bcl-2, the gene involved in human follicular lymphoma. *Proc Natl Acad Sci USA* **83**: 5214–5218.
15. Bakhshi A, Wright J, Graininger W, *et al*. (1987) Mechanism of the t(14;18) chromosomal translocation: structural analysis of both derivative 14 and 18 reciprocal partners. *Proc Natl Acad Sci USA* **84**: 2396–2400.
16. McDonnell T, Deane N, Platt F, *et al*. (1989) bcl-2-immunoglobulin transgenic mice demonstrate extended B-cell survival and follicular lymphoproliferation. *Cell* **57**: 79–88.
17. Cantor K, Blair A, Everett G, *et al*. (1992) Pesticides and other agricultural risk factors for non-Hodgkin's lymphoma among men in Iowa and Minnesota. *Cancer Res* **52**: 2447.
18. Blair A, Linos A, Stewart P, *et al*. (1993) Evaluation of risks for non-Hodgkin's lymphoma by occupation and industry exposures from a case-control study. *Am J Ind Med* **23**: 301–312.
19. Harris N, Jaffe E, Stein H, *et al*. (1994) A revised European–American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood* **84**: 1361–1392.
20. Crescenzi M, Seto M, Herzig G, Weiss P, Griffith R, Korsmeyer S (1988) Thermostable DNA polymerase chain amplification of t(14;18) chromosome breakpoints and detection of minimal residual disease. *Proc Natl Acad Sci USA* **85**: 4869–4873.
21. Liu J, Johnson R, Traweek S (1993) Rearrangement of the BCL-2 gene in follicular lymphoma. Detection by PCR in both fresh and fixed tissue samples. *Diag Mol Pathol* **2**: 241–247.
22. StataCorp (1999) Stata Statistical Software, 6.0 ed. College Station, TX: Stata Corporation.
23. Rubin D (1976) Inference and missing data. *Biometrika* **63**: 581–592.
24. Dempster A, Laird N, Rubin D (1977) Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc B* **39**: 1–38.
25. Schroeder J, Weinberg C (2001) Use of missing-data methods to correct bias and improve precision in case-control studies where cases are subtyped but subtype information is incomplete. *Am J Epidemiol* **154**: 954–962.
26. Dubin N, Pasternack B (1986) Risk assessment for case–control subgroups by polychotomous logistic regression. *Am J Epidemiol* **123**: 1101–1017.
27. Greenland S (1995) Dose–response and trend analysis in epidemiology: alternatives to categorical analysis. *Epidemiology* **6**: 356–365.
28. Schroeder J, Olshan A, Baric R, *et al*. (2001) Agricultural risk factors for t(14;18) subtypes of non-Hodgkin's lymphoma. *Epidemiology* **12**: 701–709.
29. Greenland S, Rothman K (1998) Concepts of interaction. In: Rothman K, Greenland S, eds. *Modern Epidemiology*, 2nd edn. Philadelphia: Lippincott-Raven, pp. 329–342.
30. Fifth International Workshop on Chromosomes in Leukemia–Lymphoma (1987) Correlation of chromosome abnormalities with histologic and immunologic characteristics in non-Hodgkin's lymphoma and adult T cell leukemia–lymphoma. *Blood* **70**: 1554–1564.
31. Pezzella F, Ralfkiaer E, Gatter K, Mason D (1990) The 14;18 translocation in European cases of follicular lymphoma: compar-

ison of Southern blotting and the polymerase chain reaction. *Br J Haematol* **76**: 58–64.
32. Pottern LM, Linet M, Blair A, *et al.* (1991) Familial cancers associated with subtypes of leukemia and non-Hodgkin's lymphoma. *Leuk Res* **15**: 305–314.
33. Cantor K, Blair A, Everett G, *et al.* (1988) Hair dye use and risk of leukemia and lymphoma. *Am J Public Health* **78**: 570–571.
34. Paffenbarger R, Wing A, Hyde R (1978) Characteristics in youth predictive of adult-onset malignant lymphomas, melanomas, and leukemias: brief communication. *J Natl Cancer Inst* **60**: 89–92.
35. Linet MS, McLaughlin JK, Hsing AW, *et al.* (1992) Is cigarette smoking a risk factor for non-Hodgkin's lymphoma or multiple myeloma? Results from the Lutheran Brotherhood Cohort Study. *Leuk Res* **16**: 621–624.
36. Franceschi A, Serraino D, Bidoli E, *et al.* (1989) The epidemiology of non-Hodgkin's lymphoma in the north-east of Italy: a hospital-based case–control study. *Leuk Res* **13**: 465–472.
37. Parker AS, Cerhan JR, Dick F, *et al.* (2000) Smoking and risk of non-Hodgkin lymphoma subtypes in a cohort of older women. *Leuk Lymphoma* **37**: 341–349.
38. Bell D, Liu Y, Cortopassi G (1995) Occurrence of bcl-2 oncogene translocation with increased frequency in the peripheral blood of heavy smokers. *J Natl Cancer Inst* **87**: 223–224.
39. Gordis L (1982) Should dead cases be matched to dead controls? *Am J Epidemiol* **115**: 1–5.
40. McLaughlin J, Blot W, Mehl E, Mandel J (1985) Problems in the use of dead controls in case-control studies. I. General results. *Am J Epidemiol* **121**: 131–139.
41. Centers for Disease Control and Prevention (2000) Youth Risk Behavior Surveillance–United States, 1999. *MMWR* **49**(SS-5).
42. Zahm SH, Weisenburger D, Babbitt P, Saal R, Vaught J, Blair A (1992) Use of hair coloring products and the risk of lymphoma, multiple myeloma, and chronic lymphocytic leukemia. *Am J Public Health* **82**: 990–997.
43. Altekruse SF, Henley SJ, Thun MJ (1999) Deaths from hematopoietic and other cancers in relation to permanent hair dye use in a large prospective study (United States). *Cancer Causes Control* **10**: 617–625.
44. Grodstein F, Hennekens C, Colditz G, Hunter D, Stampfer M (1994) A prospective study of permanent hair dye use and hematopoeitic cancer. *J Natl Cancer Inst* **86**: 1466–1470.
45. La Vecchia C, Tavani A (1995) Epidemiological evidence on hair dyes and the risk of cancer in humans. *Eur J Cancer Prevention* **4**: 31–43.
46. Linet MS, Pottern L (1992) Familial aggregation of hematopoietic malignancies and risk of non-Hodgkin's lymphoma. *Cancer Res* **52s**: 5468s–5473s.
47. Filipovich A, Mathur A, Kamat D, Shapiro R (1992) Primary immunodeficiencies: genetic risk factors for lymphoma. *Cancer Res* **52s**: 5645s–5467s.
48. Levine A (1992) Acquired immunodeficiency syndrome-related lymphoma. *Blood* **80**: 8–20.
49. Ho M, Miller G, Atchison RW, *et al.* (1985) Epstein–Barr virus infections and DNA hybridization studies in posttransplantation lymphoma and lymphoproliferative lesions: the role of primary infection. *J Infect Dis* **152**: 876–886.
50. Henderson S, Rowe M, Gregory C, *et al.* (1991) Induction of bcl-2 expression by Epstein–Barr virus latent membrane protein 1 protects infected B cells from programmed cell death. *Cell* **65**: 1107–1115.
51. Cantor K (1982) Farming and mortality from non-Hodgkin's lymphoma: a case–control study. *Int J Cancer* **29**: 239–247.
52. Biggar R, Melbye M (1996) Marital status in relation to Kaposi's sarcoma, non-Hodgkin's lymphoma, and anal cancer in the pre-AIDS era. *J Acquired Immune Def Syndromes Human Retrovirol* **11**: 178–182.
53. Nelson RA, Levine AM, Marks G, Bernstein L (1997) Alcohol, tobacco, and recreational drug use and the risk of non-Hodgkin's lymphoma. *Br J Cancer* **76**: 1532–1537.
54. Gail M, Pluda J, Rabkin C, *et al.* (1991) Projections of the incidence of non-Hodgkin's lymphoma related to acquired immunodeficiency syndrome. *J Natl Cancer Inst* **83**: 695–701.
55. Harnly M, Swan S, Holly E, Kelter A, Padian N (1988) Temporal trends in the incidence of non-Hodgkin's lymphoma and selected malignancies in a population with a high incidence of acquired immunodeficiency syndrome (AIDS). *Am J Epidemiol* **128**: 261–267.
56. Horsman D, Gascoyne R, Coupland R, Coldman A, Adomat S (1995) Comparison of cytogenetic analysis Southern analysis, and polymerase chain reaction for the detection of t(14;18) in follicular lymphoma. *Am J Clin Pathol* **103**: 472–478.
57. Turner GE, Ross F, Krajewski A (1995) Detection of t(14;18) in British follicular lymphoma using cytogenetics, Southern blotting, and the polymerase chain reaction. *Br J Haematol* **89**: 223–225.
58. Pezzella F, Jones M, Ralfkiaer E, Ersboll J, Gatter K, Mason D (1992) Evaluation of bcl-2 protein expression and 14;18 translocation as prognostic markers in follicular lymphoma. *Br J Cancer* **65**: 87–89.
59. Galoin S, Al Saati T, Schlaifer D, Huynh A, Attal M, Delsol G (1996) Oligonucleotide clonospecific probes directed against the functional sequence of t(14;18): a new tool for the assessment of minimal residual disease in follicular lymphomas. *Br J Hematol* **94**: 676–684.
60. Miettinen M, Lasota J (1997) Polymerase chain reaction based gene rearrangement studies in the diagnosis of follicular lymphoma–performance in formaldehyde-fixed tissue and application in clinical problem cases. *Pathol Res Pract* **193**: 9–19.
61. Corbally N, Grogan L, Devan P, Carney D (1992) The detection of specific gene rearrangements in non-Hodgkin's lymphoma using the polymerase chain reaction. *Br J Cancer* **66**: 805–809.
62. Clark HM, Jones DB, Wright DH (1992) Cytogenetic and molecular studies of t(14;18) and t(14;19) in nodal and extra-nodal B-cell lymphoma. *J Pathol* **166**: 129–137.
63. Lee K (1993) Investigation of bcl-2 rearrangements in a United Kingdom series of low and high grade non-Hodgkin's lymphomas. *Leuk Lymphoma* **11**: 91.
64. Levine E, Arthur D, Frizzera G, Peterson B, Hurd D, Bloomfield C (1985) There are differences in cytogenetic abnormalities among histologic subtypes of the non-Hodgkin's lymphomas. *Blood* **66**: 1414–1422.
65. Yunis J, Frizzera G, Oken M, McKenna J, Theologides A, Arnesen M (1987) Multiple recurrent genomic defects in follicular lymphoma. *N Engl J Med* **316**: 79–84.
66. Blair A, Zahm S (1993) Patterns of pesticide use among farmers: implications for epidemiologic research. *Epidemiology* **4**: 55–62.
67. Blair A, Zahm SH (1990) Methodologic issues in exposure assessment for case–control studies of cancer and herbicides. *Am J Ind Med* **18**: 285–293.